Standard Update

Better Data Through Standards

*The **Standard Update**, EPA s newsletter about data standards, contains information you can read and reuse whatever your interest in information management may be. For more information on any newsletter topic, see the Environmental Data Registry (EDR) Website at **http://www.epa.gov/edr/**. The EDR Website is your source for **Standard Update** back issues and a vehicle for reader comments.*

Council Charts Course for Data Standards Development

At its May 2000 meeting in Chicago, Illinois, the Environmental Data Standards Council selected four new standards–**permitting, geospatial referencing, enforcement/compliance, and tribal identifiers**–for development. The Council based their selection on recommendations from the Data Standards Priorities Workgroup, which was commissioned by the Council to develop a proposal for future data standards development priorities.

The Environmental Data Standards Council, chartered by the State/EPA Information Management Workgroup (an ongoing partnership between EPA and the Environmental Council of the States), is committed to developing and adopting data standards for environmental information exchange among EPA, states, tribes, and other parties.

The Council also endorsed the Date, Standard Industrial Classification (SIC) Code/North American Industry Classification System (NAICS), Facility Identification, and Latitude/Longitude standards, which were adopted by EPA.

Biology and Chemistry Draft Standards Business Rules Move Step Closer to Completion

In June 2000, the Data Standards Business Rules Workgroups for Biological Taxonomy and Chemical Identification completed consensus drafts on their standard business rules. These standards will be presented to EPA's Quality and Information Council (QIC) in September. The Environmental Data Standards Council will also review the draft standards business rules at their August 2000 meeting.

Business rules for EPA data standards describe what the standard is and how it will be used in conducting Agency business.

Traditionally, Agency programs have recorded information about biological organisms and chemicals of interest with a variety of names, making relating information from several sources among programmatic disciplines difficult. When the Agency biological nomenclature and taxonomy standard and the chemical identification standard are implemented, there will be more reliable exchange and aggregation of information from various programs and databases.

States and tribes are encouraged to participate in the Business Rules development and are given the opportunity to discuss and vote on the rules. States and tribes are also encouraged to adopt EPA's data standards.

Standard Tribal Names and Codes Now Available in the Environmental Data Registry

What do *Cocopah*, *Havasupai*, *Makah*, and *Yomba Shoshone* have to do with environmental metadata? They are tribal names–just part of the metadata that EPA collects and stores in the Environmental Data Registry (EDR).

EDR recently published a significant update to the Tribal Information group, consisting of tribal names, codes, etc., that includes information from the Bureau of Indian Affairs (BIA), the Indian Health Service (IHS), the National Pollutant Discharge Elimination System (NPDES), the Trust Asset and Accounting Management System (TAAMS), and the United States Census Bureau (USBC). As a result of EPA's normal data collection activities, many EPA databases store tribal names and tribal codes. EPA recently announced that it intends to develop a Tribal Identifiers data standard to promote the use of consistent tribal names and code sets across these databases.

So, What Is the EDR?

The EDR is a Web-enabled Oracle database designed to support data standards, promote data documentation, and foster information integration and sharing across the Agency and with states, tribes, and other Agency partners. The EDR does not contain the actual data from Agency information systems, but rather the information about the data–the metadata that enables a user to better understand, access, and share systems information.

The EDR serves a host of functions. It serves:

- As a directory of EPA's major information collections.
- As a dictionary of standard data elements with name, definition, and format information for system developers to use.
- As the backbone to the Agency data standards development process by providing a forum for communication with partners.
- As a data management tool, it enables users to inventory the content of the information managed by the Agency and to reduce redundancies.

The EDR organizes data elements from across the Agency by application system and by subject matter group. The data elements are uniquely defined and identified so that information can be shared throughout the Agency and integrated across Agency systems. Data elements may be associated with any kind of source including a document, standard, system, group, partner, or message set.

Data elements are the foundation of the EDR.

EPA's goal is to develop data standards that will lessen confusion caused by multiple methods of representing the same information across EPA regulations, reporting requirements, and databases. The use of data standards allows both program managers and the public greater and more rapid access to cross-Agency data. Standards help Agency managers and the public conduct more timely and accurate assessments of environmental performance, enable better use of resources, and improve data integrity. When standards and business rules become final, the EDR also provides a mechanism for workgroup and public comment through discussion databases.

For additional information on Agency data standards activities, contact Lyn Burger, Chief, Data Standards Branch, Collection Strategies Division, Office of Information Collection, 202-260-5409 or Burger.Lyn@epa.gov.